

Average output entropy for quantum channels

Christopher King¹ and David K. Moser^{1,2}

1: Department of Mathematics

2: Department of Physics

Northeastern University

Boston MA 02115

January 20, 2013

Abstract

We study the regularized average Renyi output entropy $\overline{S}_r^{\text{reg}}$ of quantum channels. This quantity gives information about the average noisiness of the channel output arising from a typical, highly entangled input state in the limit of infinite dimensions. We find a closed expression for β_r^{reg} , a quantity which we conjecture to be equal to $\overline{S}_r^{\text{reg}}$. We find an explicit form for β_r^{reg} for some entanglement-breaking channels, and also for the qubit depolarizing channel Δ_λ as a function of the parameter λ . We prove equality of the two quantities in some cases, in particular we conclude that for Δ_λ both are non-analytic functions of the variable λ .

Contents

1	Introduction	2
2	Main result	6
2.1	Notation	6
2.2	Evaluating trace moments	7
2.3	Product channels	9
2.4	Average moments of $\mathcal{A}^{\otimes n}$	9
2.5	Entanglement breaking channels	11
2.6	The qubit depolarizing channel	15
2.6.1	Evaluating $\mathcal{Q}_{\Delta_\lambda(\alpha)}$	15
2.6.2	Regularized output entropy of $\Delta_\lambda^{\otimes n}$	16
2.6.3	Output entropy of random sequences	20
3	Proof of lemmas	21
3.1	Maximal \mathcal{Q} for Δ_λ	21
3.2	Bound on Lipschitz constant	25

4	Other results	28
4.1	$\mathcal{Q}((1\dots r))$ for a more general qubit channel	28
4.2	$\mathcal{Q}((1\dots r))$ of Δ_λ for any dimension d	29
5	Conclusion	30
6	Acknowledgements	30
A	\mathcal{Q} sum diagrams	30

1 Introduction

The noisiness of a quantum channel is closely related to its ability to transfer information, and is reflected in the values of the various channel capacities. Much work has been done on understanding the capacities, for example [1] provides a recent survey. These capacities are sometimes difficult to analyze directly, for example the Holevo capacity is computed using multiple output states. Accordingly other more mathematically tractable quantities have been used to measure the amount of noise introduced by the channel. One example is the minimal output Renyi entropy [2] of a channel \mathcal{A} , defined for $r \geq 1$:

$$S_{r,\min}(\mathcal{A}) = \min_{|\phi\rangle} \frac{1}{1-r} \log \text{Tr} (\mathcal{A}(|\phi\rangle\langle\phi|))^r$$

At $r = 1$ this yields the minimal output von Neumann entropy, which has a close connection to the classical capacity of the channel [3]. The entropies for $r > 1$ also provide useful properties of the channel, and in some cases are easier to analyze and compute. The famous additivity conjecture concerns the regularized version of this quantity, which is defined as

$$S_{r,\min}^{\text{reg}}(\mathcal{A}) = \lim_{n \rightarrow \infty} \frac{1}{n} S_{r,\min}(\mathcal{A}^{\otimes n})$$

(the existence of the limit is an easy consequence of the sub-additivity bound $S_{r,\min}(\mathcal{A} \otimes \mathcal{B}) \leq S_{r,\min}(\mathcal{A}) + S_{r,\min}(\mathcal{B})$). While the inequality $S_{r,\min}^{\text{reg}}(\mathcal{A}) \leq S_{r,\min}(\mathcal{A})$ is always true, it was an open question for several years whether equality holds. It is now known that equality does not hold in general [4, 5], so this raises the interesting question of determining $S_{r,\min}^{\text{reg}}(\mathcal{A})$. Except for those channels where additivity does hold, the value of this regularized quantity is unknown. For channels with non-additive Holevo capacity the classical capacity is also defined by such a regularized quantity, so it is an important problem to find new ways to calculate these regularized limits. In a sense we follow a strategy opposite to the random channel methods used to disprove the additivity conjecture; our high-dimensional channels are products of fixed channels, and thus are constructed explicitly.

In the hopes of finding some new insights into these regularized channel properties we consider a related quantity which also measures the noisiness of the

channel, namely the average output Renyi entropy. This measures the entropy of the channel output for typical input states, rather than the smallest value which is used to compute the minimal output entropy. For a finite-dimensional channel this is defined for all $r \geq 1$ by

$$\overline{S}_r(\mathcal{A}) = \mathbb{E} \left[\frac{1}{1-r} \log \text{Tr} (\mathcal{A}(|\phi\rangle\langle\phi|))^r \right]$$

where the expectation is computed using the uniform probability measure on the set of input pure states. We also consider the related quantity

$$\beta_r(\mathcal{A}) = \frac{1}{1-r} \log \mathbb{E} [\text{Tr} (\mathcal{A}(|\phi\rangle\langle\phi|))^r]$$

Note that by Jensen's inequality

$$\overline{S}_r(\mathcal{A}) \geq \beta_r(\mathcal{A}). \quad (1)$$

These quantities can be computed (at least numerically) for any given channel. In operational terms, they describe the long-run average output Renyi entropy of the channel for a sequence of random pure input states. Loosely speaking, they measure the average noisiness of an output state from the channel.

As with the minimal Renyi entropy, we also consider the regularized versions of these quantities. However unlike the minimal Renyi entropy, the existence of these regularized limits is not obvious, so we define them conservatively using the \liminf :

$$\begin{aligned} \overline{S}_r^{\text{reg}}(\mathcal{A}) &= \liminf_{n \rightarrow \infty} \frac{1}{n} \overline{S}_r(\mathcal{A}^{\otimes n}) \\ \beta_r^{\text{reg}}(\mathcal{A}) &= \liminf_{n \rightarrow \infty} \frac{1}{n} \beta_r(\mathcal{A}^{\otimes n}) \end{aligned} \quad (2)$$

We conjecture that the two quantities in (2) are equal, however we do not yet have a proof of this for a general channel. For the specific channels we look at in more detail the quantity $\beta_r^{\text{reg}}(\mathcal{A})$ is given by the expression above with \liminf replaced by \lim . But from the closed expression for $\beta_r^{\text{reg}}(\mathcal{A})$ presented below there is a possibility for more complex limiting behavior.

For one special class of channels we can compute a simple formula for $\beta_r^{\text{reg}}(\mathcal{A})$ for integer values of r . These are a subset of the entanglement breaking (E-B) channels [6, 7], which can be written in the form $\mathcal{A}(\rho) = \sum_k \sigma_k \text{Tr}(X_k \rho)$ – for some states σ_k and with X_k a POVM – meeting the additional condition $\text{Tr}(\prod_{i=1}^r \sigma_{k_i}) \geq 0$. For $r = 2$ this includes all the E-B channels. For higher r a particular class of E-B channels that fulfill the condition are the QC channels as defined by Holevo [8], where $\sigma_k = |k\rangle\langle k|$ are pure states formed by an orthonormal basis. For unital E-B channels which satisfy the condition on the σ_k we can prove even more, namely that $\overline{S}_r^{\text{reg}}(\mathcal{A}) = \beta_r^{\text{reg}}(\mathcal{A}) = \log d$.

We consider only finite-dimensional channels with equal input and output dimensions, and we define the dimension of the channel to be this common value. The identity matrix is denoted by $\mathbb{1}$.

Theorem 1 (proof in 2.5).

- (a) Let $r \geq 2$ be an integer, and $\mathcal{A}(\rho) = \sum \sigma_k \text{Tr}(X_k \rho)$ an entanglement breaking channel satisfying the condition $\text{Tr}(\prod_{i=1}^r \sigma_{k_i}) \geq 0$ for all choices of $\{k_i\}$. Then

$$\beta_r^{\text{reg}}(\mathcal{A}) = \lim_{n \rightarrow \infty} \frac{1}{n} \beta_r(\mathcal{A}^{\otimes n}) = \frac{1}{1-r} \log \text{Tr}(\mathcal{A}(\mathbb{1}/d)^r).$$

- (b) Let \mathcal{A} be a d -dimensional unital entanglement-breaking channel, $\mathcal{A}(\mathbb{1}) = \mathbb{1}$, satisfying the condition $\text{Tr}(\prod_{i=1}^m \sigma_{k_i}) \geq 0$ for all integer $m \geq 2$. Then for all real $r \geq 1$

$$\overline{S}_r^{\text{reg}}(\mathcal{A}) = \beta_r^{\text{reg}}(\mathcal{A}) = \log d.$$

We also derive an explicit expression for $\beta_r^{\text{reg}}(\mathcal{A})$ in the general case. The statement of this result requires some additional notation. First recall the definition of the Choi-Jamiolkowski representation [9] of a channel, namely

$$\text{Choi}(\mathcal{A}) = \sum_{x,y} \mathcal{A}(|x\rangle\langle y|) \otimes |x\rangle\langle y|$$

where $|x\rangle$ and $|y\rangle$ are orthonormal bases of pure input states. Also let $\text{Sym}(r)$ denote the symmetric group on r letters. Then every element $\alpha \in \text{Sym}(r)$ defines a permutation operator on $(\mathbb{C}^d)^{\otimes r}$ by $\mathcal{R}(\alpha)(v_1 \otimes \cdots \otimes v_r) = v_{\alpha(1)} \otimes \cdots \otimes v_{\alpha(r)}$.

Definition 2. Let \mathcal{A} be a channel. For all $\alpha \in \text{Sym}(r)$ define

$$\mathcal{Q}_{\mathcal{A},r}(\alpha) = \text{Tr} [\text{Choi}(\mathcal{A})^{\otimes r} (\mathcal{R}(123 \dots r) \otimes \mathcal{R}(\alpha))] . \quad (3)$$

Furthermore, for some channels β_r^{reg} is a simple limit. One such class are the entrywise positive maps defined studied in [10].

Definition 3. A channel \mathcal{A} is called *entrywise positive* if there exist a bases for input and output space such that $\langle s | \mathcal{A}(|x\rangle\langle y|) | t \rangle \geq 0$ for all x, y, s, t .

It is clear that this definition is equivalent to $\text{Choi}(\mathcal{A})$ being entrywise positive.

Theorem 4 (proof in 2.4).

- (a) Let $r \geq 2$ be an integer and let $\mathcal{Q}_{\max} = \max_{\alpha \in \text{Sym}(r)} |\mathcal{Q}_{\mathcal{A},r}(\alpha)|$. Then

$$\beta_r^{\text{reg}}(\mathcal{A}) = \frac{r \log d - \log \mathcal{Q}_{\max}}{r-1}.$$

- (b) Let $r \geq 2$ be an integer. If the maximum \mathcal{Q}_{\max} is attained for a unique α then the \liminf in β_r^{reg} can be replaced with a regular limit:

$$\beta_r^{\text{reg}}(\mathcal{A}) = \lim_{n \rightarrow \infty} \frac{1}{n} \beta_r(\mathcal{A}^{\otimes n}).$$

(c) If the channel \mathcal{A} is entrywise positive then for all integer $r \geq 2$

$$\beta_r^{\text{reg}}(\mathcal{A}) = \lim_{n \rightarrow \infty} \frac{1}{n} \beta_r(\mathcal{A}^{\otimes n}).$$

The evaluation of \mathcal{Q}_{\max} seems to be a difficult problem in general for large values of r . However for two special permutations the quantity $\mathcal{Q}_{\mathcal{A},r}(\alpha)$ can be evaluated easily, namely the identity permutation and the full cycle:

$$\mathcal{Q}_{\mathcal{A},r}(\text{id}) = \text{Tr } \mathcal{A}(\mathbb{1})^r, \quad \mathcal{Q}_{\mathcal{A},r}(123 \dots r) = \text{Tr } (\text{Choi}(\mathcal{A})^r).$$

Thus for $r = 2$ the result can be stated more explicitly as follows.

Corollary 5. *Let \mathcal{A} be a d -dimensional channel, then*

$$\beta_2^{\text{reg}}(\mathcal{A}) = 2 \log d - \log \max[\text{Tr } \mathcal{A}(\mathbb{1})^2, \text{Tr } (\text{Choi}(\mathcal{A})^2)]$$

To make further progress we now focus on one of the simplest cases, namely the qubit depolarizing channel Δ_λ [11], where we are able to prove a number of additional results. In particular using concentration of measure arguments [12, 13] we compute the regularized quantity $\bar{S}_r^{\text{reg}}(\Delta_\lambda)$ – which we call regularized output entropy in the remainder of the paper – for integer values of r , and for a range of values of λ . One interesting consequence is that this quantity is a non-analytic function of the depolarizing parameter λ . Recall the definition of this channel:

$$\Delta_\lambda(\rho) = \lambda \rho + \frac{1-\lambda}{2} \mathbb{1}$$

The channel is completely positive for $-1/3 \leq \lambda \leq 1$ and is entanglement breaking for $-1/3 \leq \lambda \leq 1/3$.

Theorem 6 (proof in 2.6.2).

(a) For all $r \in \mathbb{N}$, $r \geq 2$, and $\lambda \in [0, 1]$,

$$\beta_r^{\text{reg}}(\Delta_\lambda) = \lim_{n \rightarrow \infty} \frac{1}{n} \beta_r(\Delta_\lambda^{\otimes n}) = \min \left\{ 1, \frac{2r - \log[(1+3\lambda)^r + 3(1-\lambda)^r]}{r-1} \right\}$$

in particular

$$\beta_2^{\text{reg}}(\Delta_\lambda) = \begin{cases} 1 & \lambda \leq 1/\sqrt{3} \\ 2 - \log(1+3\lambda^2) & \lambda > 1/\sqrt{3} \end{cases},$$

$$\beta_\infty^{\text{reg}}(\Delta_\lambda) = \begin{cases} 1 & \lambda \leq 1/3 \\ 2 - \log(1+3\lambda) & \lambda > 1/3 \end{cases},$$

where $\beta_\infty^{\text{reg}}(\Delta_\lambda) = \lim_{r \rightarrow \infty} \beta_r^{\text{reg}}(\Delta_\lambda)$.

(b) For all $r \in \mathbb{N}$, $r \geq 2$, and $\lambda \in J_r$,

$$\overline{S}_r^{\text{reg}}(\Delta_\lambda) = \beta_r^{\text{reg}}(\Delta_\lambda),$$

where $J_r = [0, c_r] \cup [d_r, 1] \subset [0, 1]$ for some $0 < c_r < d_r < 1$ (see Table 1 in 2.6.2).

From the explicit form given in (a) it is clear that $\beta_r^{\text{reg}}(\Delta_\lambda)$ does not have a continuous first derivative for some $\lambda_r \in [1/3, 1/\sqrt{3}]$, in particular it is non-analytic. Because $\overline{S}_r^{\text{reg}}(\Delta_\lambda)$ is defined as the liminf of a series upper bounded by 1 it is well defined for all $0 \leq \lambda \leq 1$. Furthermore we know from (b) that it is equal to 1 for some range $\lambda \in [0, c_r]$ but at $\lambda = 1$ its value is 0. Therefore $\overline{S}_r^{\text{reg}}(\Delta_\lambda)$ has least one non-analytic point somewhere in the range $\lambda \in [c_r, d_r]$.

It is tempting to associate this non-analyticity with a transition between distinct phases of the model, but the operational meaning of this is unclear at the moment. We conjecture that the quantities $\overline{S}_r^{\text{reg}}(\Delta_\lambda)$ and $\beta_r^{\text{reg}}(\Delta_\lambda)$ are in fact equal for all λ , and for all $r \geq 1$. It is noteworthy that the channel Δ_λ is entanglement-breaking at and below the value $\lambda = 1/3$.

Counterexamples to the additivity conjecture have been found so far by using randomization techniques [14, 15, 16, 5]. This has led to an understanding of the behavior of a typical high-dimensional channel, at least insofar as it affects the minimal output Renyi entropy, by proving the generic existence of channels all of whose output states have high entropy. Here we look from a different point of view, by considering the properties of a typical output state for a product of many copies of a channel. Open questions remain, for example the amount of entanglement in a typical output state. We note that the questions addressed here have a different flavor from arguments based on locality, since here the system is fully entangled across all copies.

2 Main result

2.1 Notation

We work with a general channel \mathcal{A} and its tensor product $\mathcal{C} = \mathcal{A}^{\otimes n}$. The dimension of \mathcal{A} is

$$d = \dim \mathcal{A}$$

To achieve the results for $\overline{S}_r^{\text{reg}}$ and β_r^{reg} in Theorems 1, 4 and 6 we first find a closed expression for the average moments for integer $r \geq 2$

$$M_r(\mathcal{C}) = \mathbb{E} [\text{Tr} (\mathcal{C}(|\phi\rangle\langle\phi|)^r)], \quad (4)$$

where averaging is over random pure input states $|\phi\rangle = U|0\rangle$ with U distributed according to the Haar measure on $SU(d^n)$. We then use the relation

$$\beta_r^{\text{reg}}(\mathcal{A}) = \liminf_{n \rightarrow \infty} \frac{1}{n(1-r)} \log M_r(\mathcal{A}^{\otimes n})$$

2.2 Evaluating trace moments

We rewrite the trace moment (4) by inserting four complete sums that run over the entire input space

$$\begin{aligned}
& \mathbb{E} [\text{Tr} (\mathcal{C}(|\phi\rangle\langle\phi|)^r)] \\
&= \mathbb{E} \left[\text{Tr} \left(\sum_{a,b,x,y} |a\rangle\langle a| \mathcal{C}(|x\rangle\langle x| \phi |y\rangle\langle y|) |b\rangle\langle b| \right)^r \right] \\
&= \mathbb{E} \left[\sum_{\substack{\{a_i, x_i, y_i\} \\ i=1 \dots r}} \prod_{i=1}^r \mathcal{C}_{a_i x_i y_i a_{i+1}} \langle x_i | \phi \rangle \langle \phi | y_i \rangle \right] \\
&= \sum_{\substack{\{a_i, x_i, y_i\} \\ i=1 \dots r}} \prod_{i=1}^r \mathcal{C}_{a_i x_i y_i a_{i+1}} \mathbb{E} \left[\prod_{j=1}^r \langle x_j | \phi \rangle \langle \phi | y_j \rangle \right] \tag{5}
\end{aligned}$$

with the identification $a_{r+1} \equiv a_1$, and with

$$\mathcal{C}_{axyb} = \langle a | \mathcal{C}(|x\rangle\langle y|) | b \rangle$$

the matrix elements of the channel. The expectation value in (5)

$$\mathbb{E} \left[\prod_{j=1}^r \langle x_j | \phi \rangle \langle \phi | y_j \rangle \right] = \mathbb{E} \left[\prod_{j=1}^r \langle x_j | U | 0 \rangle \langle 0 | U^* | y_j \rangle \right] \tag{6}$$

of products of matrix elements of unitaries distributed according to the Haar measure may be calculated using Weingarten calculus [17]. The Weingarten function $\text{Wg} : \mathbb{N} \times \text{Sym}(r) \rightarrow \mathbb{R}$ maps pairs of dimension k and elements of the symmetric group $\text{Sym}(r)$ into the reals. The general expression is

$$\begin{aligned}
& \mathbb{E} [U_{i_1 j_1} \dots U_{i_r j_r} \overline{U_{i'_1 j'_1}} \dots \overline{U_{i'_r j'_r}}] \\
&= \sum_{\alpha, \beta \in \text{Sym}(r)} \delta_{i_{\alpha(1)} i'_1} \dots \delta_{i_{\alpha(r)} i'_r} \delta_{j_{\beta(1)} j'_1} \dots \delta_{j_{\beta(r)} j'_r} \text{Wg}(k, \beta^{-1} \alpha) .
\end{aligned}$$

Thus (6) simplifies to

$$\mathbb{E} \left[\prod_{j=1}^r \langle x_j | U | 0 \rangle \langle 0 | U^* | y_j \rangle \right] = \sum_{\alpha, \beta \in \text{Sym}(r)} \delta_{x_{\alpha(1)} y_1} \dots \delta_{x_{\alpha(r)} y_r} \text{Wg}(k, \beta \alpha^{-1}) \tag{7}$$

$$= \sum_{\alpha \in \text{Sym}(r)} \delta_{x_{\alpha(1)} y_1} \dots \delta_{x_{\alpha(r)} y_r} C_{k, r} \tag{8}$$

where k is the input dimension for \mathcal{C} , and in the last step the Weingarten function is summed over all permutations $\gamma = \beta \alpha^{-1}$. This sum can be evaluated explicitly, as was shown for example in [18]:

$$C_{k, r} = \sum_{\gamma \in \text{Sym}(r)} \text{Wg}(k, \gamma) = \prod_{j=0}^{r-1} \frac{1}{k+j} . \tag{9}$$

We plug the evaluated expectation value (8) in our original expression (5) and get

$$\begin{aligned} \sum_{\substack{\{a_i, x_i, y_i\} \\ i=1 \dots r}} \left(\prod_{i=1}^r c_{a_i x_i y_i a_{i+1}} \cdot \sum_{\alpha \in \text{Sym}(r)} \prod_{j=1}^r \delta_{x_{\alpha(j)} y_j} \cdot C_{k, r} \right) \\ = C_{k, r} \sum_{\substack{\{a_i, x_i\} \\ i=1 \dots r}} \sum_{\alpha \in \text{Sym}(r)} \prod_{i=1}^r c_{a_i x_i x_{\alpha(i)} a_{i+1}}. \end{aligned} \quad (10)$$

Now define

$$\mathcal{Q}_{\mathcal{C}, r}(\alpha) = \sum_{\substack{\{a_i, x_i\} \\ i=1 \dots r}} \prod_{i=1}^r c_{a_i x_i x_{\alpha(i)} a_{i+1}} = \sum_{\substack{\{x_i\} \\ i=1 \dots r}} \text{Tr} \prod_{i=1}^r c_{x_i x_{\alpha(i)}}, \quad (11)$$

where the matrices \mathcal{C}_{xy} have entries $(\mathcal{C}_{xy})_{ab} = \mathcal{C}_{axyb}$ or more simply $\mathcal{C}_{xy} = \mathcal{C}(|x\rangle\langle y|)$. If it is clear from context we may omit one or both subscripts $\mathcal{Q}(\alpha) = \mathcal{Q}_{\mathcal{C}}(\alpha) = \mathcal{Q}_{\mathcal{C}, r}(\alpha)$. These are the terms we will analyze to a great length in the rest of the work.

The \mathcal{Q} defined in this way is identical to the one from Definition 2. This can be seen from the following calculation

$$\begin{aligned} & \text{Tr} (C_{\text{hoi}}(\mathcal{C})^{\otimes r} (\mathcal{R}(123 \dots r) \otimes \mathcal{R}(\alpha))) \\ &= \sum_{\{a, x\}} \langle a_1, x_1 | \otimes \dots \otimes \langle a_r, x_r | C_{\text{hoi}}(\mathcal{C})^{\otimes r} | a_2, x_{\alpha(1)} \rangle \otimes \dots \otimes | a_1, x_{\alpha(r)} \rangle \\ &= \sum_{\{a, x\}} \langle a_1 | \mathcal{C}(|x_1\rangle\langle x_{\alpha(1)}|) | a_2 \rangle \langle a_2 | \mathcal{C}(|x_2\rangle\langle x_{\alpha(2)}|) | a_3 \rangle \dots \\ &= \sum_{\substack{\{a_i, x_i\} \\ i=1 \dots r}} \prod_{i=1}^r c_{a_i x_i x_{\alpha(i)} a_{i+1}} \\ &= \mathcal{Q}_{\mathcal{C}, r}(\alpha). \end{aligned}$$

In terms of the $\mathcal{Q}_{\mathcal{C}}$ we get our final working expression for the average moments

$$M_r(\mathcal{C}) = C_{k, r} \sum_{\alpha \in \text{Sym}(r)} \mathcal{Q}_{\mathcal{C}}(\alpha).$$

In general $\mathcal{Q}_{\mathcal{C}}(\alpha)$ could have complex values. However in some cases it can be shown to be real. In particular for the depolarizing channel it follows directly from $\mathcal{C}_{axyb} \geq 0$ that $\mathcal{Q}_{\mathcal{C}}(\alpha)$ is positive.

2.3 Product channels

In the case where \mathcal{C} is a tensor product $\mathcal{C} = \mathcal{D} \otimes \mathcal{E}$ we work in the product base $|x\rangle = |x'x''\rangle$. Now the tensor and channel application are interchangeable

$$\begin{aligned}\mathcal{D} \otimes \mathcal{E}_{xy} &= \mathcal{D} \otimes \mathcal{E}(|x'x''\rangle\langle y'y''|) \\ &= \mathcal{D}(|x'\rangle\langle y'|) \otimes \mathcal{E}(|x''\rangle\langle y''|) \\ &= \mathcal{D}_{x'y'} \otimes \mathcal{E}_{x''y''}.\end{aligned}$$

And therefore, the $\mathcal{Q}_{\mathcal{D} \otimes \mathcal{E}}(\alpha)$ factors

$$\begin{aligned}\mathcal{Q}_{\mathcal{D} \otimes \mathcal{E}}(\alpha) &= \sum_{\substack{\{x_i\} \\ i=1\dots r}} \text{Tr} \prod_{i=1}^r \mathcal{D} \otimes \mathcal{E}_{x_i x_{\alpha(i)}} \\ &= \sum_{\substack{\{x'_i, x''_i\} \\ i=1\dots r}} \text{Tr} \prod_{i=1}^r \mathcal{D}_{x'_i x'_{\alpha(i)}} \otimes \mathcal{E}_{x''_i x''_{\alpha(i)}} \\ &= \sum_{\substack{\{x'_i\} \\ i=1\dots r}} \text{Tr} \prod_{i=1}^r \mathcal{D}_{x'_i x'_{\alpha(i)}} \sum_{\substack{\{x''_i\} \\ i=1\dots r}} \text{Tr} \prod_{i=1}^r \mathcal{E}_{x''_i x''_{\alpha(i)}} \\ &= \mathcal{Q}_{\mathcal{D}}(\alpha) \mathcal{Q}_{\mathcal{E}}(\alpha).\end{aligned}\tag{12}$$

2.4 Average moments of $\mathcal{A}^{\otimes n}$

If we set $\mathcal{C} = \mathcal{A}^{\otimes n}$ the average moment factors as above, the dimension is $k = d^n$ and according to (12) we get

$$M_r(\mathcal{A}^{\otimes n}) = C_{d^n, r} \sum_{\alpha \in \text{Sym}(r)} \mathcal{Q}_{\mathcal{A}, r}(\alpha)^n.\tag{13}$$

When the meaning is clear from the context we suppress the index in $\mathcal{Q}_{\mathcal{A}}$. The limiting behavior of this sum is relatively simple and determines the quantity $\beta_r^{\text{reg}}(\mathcal{A})$ as described in the following theorem.

Theorem 4. (a) Let $r \geq 2$ be an integer and let $\mathcal{Q}_{\max} = \max_{\alpha \in \text{Sym}(r)} |\mathcal{Q}_{\mathcal{A}, r}(\alpha)|$. Then

$$\beta_r^{\text{reg}}(\mathcal{A}) = \frac{r \log d - \log \mathcal{Q}_{\max}}{r - 1}.$$

(b) Let $r \geq 2$ be an integer. If the maximum \mathcal{Q}_{\max} is attained for a unique α then the \liminf in β_r^{reg} can be replaced with a regular limit:

$$\beta_r^{\text{reg}}(\mathcal{A}) = \lim_{n \rightarrow \infty} \frac{1}{n} \beta_r(\mathcal{A}^{\otimes n}).$$

(c) If the channel \mathcal{A} is entrywise positive then for all integer $r \geq 2$

$$\beta_r^{\text{reg}}(\mathcal{A}) = \lim_{n \rightarrow \infty} \frac{1}{n} \beta_r(\mathcal{A}^{\otimes n}).$$

Proof. (a) From the definition of (2) and (13) we have

$$\begin{aligned}\beta_r^{\text{reg}}(\mathcal{A}) &= \liminf_{n \rightarrow \infty} \frac{1}{n(1-r)} \log M_r(\mathcal{A}^{\otimes n}) \\ &= \liminf_{n \rightarrow \infty} \frac{1}{n(1-r)} \log \left[C_{d^n, r} \sum_{\alpha} \mathcal{Q}(\alpha)^n \right]\end{aligned}$$

Note that

$$\begin{aligned}C_{d^n, r} \sum_{\alpha} \mathcal{Q}(\alpha)^n &= \left| C_{d^n, r} \sum_{\alpha} \mathcal{Q}(\alpha)^n \right| \\ &\leq C_{d^n, r} \sum_{\alpha} |\mathcal{Q}(\alpha)|^n \\ &\leq d^{-nr} \sum_{\alpha} \mathcal{Q}_{\max}^n \\ &= r! d^{-nr} \mathcal{Q}_{\max}^n\end{aligned}$$

Since $M_r(\mathcal{A}^{\otimes n}) \leq 1$ we have

$$\begin{aligned}\beta_r^{\text{reg}}(\mathcal{A}) &= \frac{1}{r-1} \liminf_{n \rightarrow \infty} \frac{1}{n} \log \left[C_{d^n, r} \sum_{\alpha} \mathcal{Q}(\alpha)^n \right]^{-1} \\ &\geq \frac{1}{r-1} \liminf_{n \rightarrow \infty} \frac{1}{n} \log (r! d^{-nr} \mathcal{Q}_{\max}^n)^{-1} \\ &= \frac{1}{r-1} \log (d^r \mathcal{Q}_{\max}^{-1}) \\ &= \frac{r \log d - \log \mathcal{Q}_{\max}}{r-1}.\end{aligned}$$

In order to prove equality, we will use the existence of a subsequence $\{n_j\}$ such that

$$\lim_{j \rightarrow \infty} \frac{1}{n_j(1-r)} \log \left[C_{d^{n_j}, r} \sum_{\alpha} \mathcal{Q}(\alpha)^{n_j} \right] \leq \frac{r \log d - \log \mathcal{Q}_{\max}}{r-1}. \quad (14)$$

To this end, let $\{\alpha_1, \dots, \alpha_N\}$ be the maximizers satisfying $|\mathcal{Q}(\alpha_i)| = \mathcal{Q}_{\max}$, so that $\mathcal{Q}(\alpha_i) = \mathcal{Q}_{\max} e^{2\pi i \gamma_i}$ for some $\gamma_i \in [0, 1)$. It is a basic result from simultaneous Diophantine approximations (using the Dirichlet box principle) that for any $\epsilon > 0$ there is an increasing sequence of positive integers n_j such that $\max_i \{n_j \gamma_i\} < \epsilon$ for all j , where $\{x\}$ denotes the distance to the closest integer. Choose $\epsilon = 1/6$, then we have

$$\begin{aligned}\left| \sum_{i=1}^N \mathcal{Q}(\alpha_i)^{n_j} \right| &= \mathcal{Q}_{\max}^{n_j} \left| \sum_{i=1}^N e^{2\pi i n_j \gamma_i} \right| \\ &\geq \mathcal{Q}_{\max}^{n_j} \cdot \frac{N}{2}.\end{aligned} \quad (15)$$

Furthermore, there is $\Theta < 1$ such that

$$|\mathcal{Q}(\alpha')| \leq \mathcal{Q}_{\max} \cdot \Theta$$

for all α' which are not maximizers. Thus

$$\begin{aligned} C_{d^{n_j}, r} \sum_{\alpha} \mathcal{Q}(\alpha)^{n_j} &= C_{d^{n_j}, r} \left| \sum_{\alpha} \mathcal{Q}(\alpha)^{n_j} \right| \\ &\geq C_{d^{n_j}, r} \mathcal{Q}_{\max}^{n_j} \cdot \frac{N}{2} - C_{d^{n_j}, r} \left| \sum_{\alpha'} \mathcal{Q}(\alpha')^{n_j} \right| \\ &\geq C_{d^{n_j}, r} \mathcal{Q}_{\max}^{n_j} \left(\frac{N}{2} - r! \Theta^{n_j} \right) \end{aligned}$$

Since $\Theta < 1$, for j sufficiently large we have $\frac{N}{2} - r! \Theta^{n_j} \geq \frac{1}{3}$, thus for j sufficiently large

$$\begin{aligned} \frac{1}{n_j(1-r)} \left[\log C_{d^{n_j}, r} \sum_{\alpha} \mathcal{Q}(\alpha)^{n_j} \right] \\ \leq \frac{1}{r-1} \frac{1}{n_j} \log \left(3 \mathcal{Q}_{\max}^{-n_j} C_{d^{n_j}, r}^{-1} \right). \end{aligned}$$

Using $\lim_{n \rightarrow \infty} (d^{nr} C_{d^n, r}) = 1$, the result follows immediately.

(b) Let $\alpha_0 \in \text{Sym}(r)$ be the unique permutation satisfying $|\mathcal{Q}(\alpha_0)| = \mathcal{Q}_{\max}$, so that $N = 1$ in the notation of (a). Then (15) is replaced by the equality

$$|\mathcal{Q}(\alpha_0)^n| = \mathcal{Q}_{\max}^n$$

which holds for every n . Thus the inequality (14) is true for every n , hence the upper and lower bound yield the existence of the limit.

(c) If \mathcal{A} is entrywise positive then every term $\mathcal{Q}(\alpha)$ is also positive. Thus the inequality (15) is replaced by

$$\left| \sum_{i=1}^N \mathcal{Q}(\alpha_i)^n \right| = N \mathcal{Q}_{\max}^n$$

and this holds for every n . Thus again (14) is true for every n , and the result follows. \square

2.5 Entanglement breaking channels

After dealing with basic facts about β_r^{reg} we turn our attention to the special case of entanglement breaking channels. In this case calculating the relevant \mathcal{Q} -terms and studying their properties is particularly easy. We restate and then prove Theorem 1 from the introduction.

Theorem 1. (a) Let $r \geq 2$ be an integer, and $\mathcal{A}(\rho) = \sum \sigma_k \text{Tr}(X_k \rho)$ an entanglement breaking channel satisfying the condition $\text{Tr}(\prod_{i=1}^r \sigma_{k_i}) \geq 0$ for all choices of $\{k_i\}$. Then

$$\beta_r^{\text{reg}}(\mathcal{A}) = \lim_{n \rightarrow \infty} \frac{1}{n} \beta_r(\mathcal{A}^{\otimes n}) = \frac{1}{1-r} \log \text{Tr}(\mathcal{A}(\mathbb{1}/d)^r).$$

(b) Let \mathcal{A} be a d -dimensional unital entanglement-breaking channel, $\mathcal{A}(\mathbb{1}) = \mathbb{1}$, satisfying the condition $\text{Tr}(\prod_{i=1}^m \sigma_{k_i}) \geq 0$ for all integer $m \geq 2$. Then for all real $r \geq 1$

$$\overline{S}_r^{\text{reg}}(\mathcal{A}) = \beta_r^{\text{reg}}(\mathcal{A}) = \log d.$$

Proof. (a) The result follows immediately from Lemma 7 and Theorem 4.

(b) For a unital entanglement breaking channel $\beta_r^{\text{reg}}(\mathcal{A})$ attains its maximal value $\log d$ which is at the same time the maximal possible value of $\overline{S}_r^{\text{reg}}(\mathcal{A})$. So, with inequality (1) it follows that $\overline{S}_r^{\text{reg}}(\mathcal{A}) = \beta_r^{\text{reg}}(\mathcal{A}) = \log d$ for all integer $r \geq 2$. In the following we extend this result to all real $r \geq 1$.

Consider the derivatives of the function $f_n(r) = \frac{1}{n} \ln \mathbb{E}[\text{Tr}(\mathcal{A}^{\otimes n}(|\phi\rangle\langle\phi|)^r)] / \ln 2$ with respect to r , and set $\rho = \mathcal{A}^{\otimes n}(|\phi\rangle\langle\phi|)$

$$\begin{aligned} f'_n(r) &= \frac{1}{n} \mathbb{E}[\text{Tr} \rho^r]^{-1} \mathbb{E}[\text{Tr}(\rho^r \ln \rho)] / \ln 2 < 0 \\ f''_n(r) &= \frac{1}{n} \mathbb{E}[\text{Tr} \rho^r]^{-1} \mathbb{E}[\text{Tr}(\rho^r (\ln \rho)^2)] / \ln 2 - \frac{1}{n} \mathbb{E}[\text{Tr} \rho^r]^{-2} \mathbb{E}[\text{Tr}(\rho^r \ln \rho)]^2 / \ln 2 > 0 \end{aligned} \quad (16)$$

To prove the second inequality we use two applications of the Cauchy-Schwarz inequality. First $|\text{Tr}(AB)| \leq \text{Tr}(A^2)^{1/2} \text{Tr}(B^2)^{1/2}$ with $A = \rho^{r/2}$ and $B = \rho^{r/2} \ln \rho$, and then $\mathbb{E}[XY]^2 \leq \mathbb{E}[X^2] \mathbb{E}[Y^2]$ to deduce

$$\mathbb{E}[\text{Tr}(\rho^r \ln \rho)]^2 \leq \mathbb{E}[(\text{Tr}(\rho^r (\ln \rho)^2))^{1/2} (\text{Tr} \rho^r)^{1/2}]^2 \leq \mathbb{E}[\text{Tr}(\rho^r (\ln \rho)^2)] \mathbb{E}[\text{Tr} \rho^r]$$

Therefore, the function $f_n(r)$ is convex in r for $r > 0$. We also know that $f_n(r) \geq (1-r) \log d$ for any real $r \geq 1$ and $\lim_{n \rightarrow \infty} f_n(r) = (1-r) \log d$ for any integer $r \geq 1$. Therefore, for any integer $r_0 \geq 1$ and $r \in [r_0, r_0 + 1]$ we have upper and lower bounds

$$f_n(r_0)(r_0 + 1 - r) + f_n(r_0 + 1)(r - r_0) \geq f_n(r) \geq (1 - r) \log d. \quad (17)$$

Thus we have $\lim_{n \rightarrow \infty} f_n(r) = (1-r) \log d$ for all real $r \geq 1$. Dividing by $1-r$ gives $\lim_{n \rightarrow \infty} \frac{1}{n} S_r(\mathcal{A}^{\otimes n}) = \log d$ the desired equality for all $r > 1$.

Finally, for $r = 1$ the Renyi entropy is defined as the Neumann entropy, which equals the limit $\frac{1}{n} S_1(\mathcal{A}^{\otimes n}) = \lim_{r \rightarrow 1} \frac{f_n(r)}{1-r} = -f'_n(1)$. Again, using the bound (17) with $r_0 = 1$ (and noting that $f_n(1) = 0$) we get $-f'_n(2)(1-r) \geq f_n(r) \geq (1-r) \log d$ and so

$$-f'_n(2) \leq \frac{f_n(r)}{1-r} \leq \log d$$

which implies $-f_n(2) \leq \frac{1}{n}S_1(\mathcal{A}^{\otimes n}) \leq \log d$. But as $\lim_{n \rightarrow \infty} -f_n(2) = \log d$ both sides of the bound become equal and we have $\lim_{n \rightarrow \infty} \frac{1}{n}S_1(\mathcal{A}^{\otimes n}) = \log d$ as well. \square

Lemma 7. *Let \mathcal{A} be an entanglement breaking channel with*

$$\mathrm{Tr} \left(\prod_{i=1}^r \sigma_{k_i} \right) \geq 0 \quad (18)$$

for all choices of $\{k_i\}$, then $\alpha = \mathrm{id}$ is the unique maximum of $\mathcal{Q}_{\mathcal{A}, r}$ so $\mathcal{Q}_{\max} = \mathcal{Q}_{\mathcal{A}, r}(\mathrm{id}) = \mathrm{Tr}(\mathcal{A}(\mathbb{1})^r)$.

Proof. Entanglement breaking channels are of the form $\mathcal{A}(\rho) = \sum \sigma_k \mathrm{Tr}(X_k \rho)$ where the σ_k are density matrices and the X_k constitute a POVM. Therefore we calculate

$$\begin{aligned} |\mathcal{Q}(\alpha)| &= \left| \sum_{\substack{\{x_i\} \\ i=1 \dots r}} \mathrm{Tr} \prod_{i=1}^r \mathcal{A}_{x_i x_{\alpha(i)}} \right| \\ &= \left| \sum_{\substack{\{x_i\} \\ i=1 \dots r}} \mathrm{Tr} \left(\prod_{i=1}^r \sum_k \sigma_k \mathrm{Tr}(X_k |x_i\rangle \langle x_{\alpha(i)}|) \right) \right| \\ &= \left| \sum_{\substack{\{x_i, k_i\} \\ i=1 \dots r}} \mathrm{Tr} \left(\prod_{i=1}^r \sigma_{k_i} \right) \prod_{i=1}^r \mathrm{Tr}(X_{k_i} |x_i\rangle \langle x_{\alpha(i)}|) \right| \\ &\leq \sum_{\substack{\{k_i\} \\ i=1 \dots r}} \mathrm{Tr} \left(\prod_{i=1}^r \sigma_{k_i} \right) \left| \sum_{\substack{\{x_i\} \\ i=1 \dots r}} \prod_{i=1}^r \langle x_{\alpha(i)} | X_{k_i} | x_i \rangle \right| \end{aligned} \quad (19)$$

where condition (18) is used in the last step. The term on the right side of the last line can be rewritten as a product of traces

$$\begin{aligned} \sum_{\substack{\{x_i\} \\ i=1 \dots r}} \prod_{i=1}^r \langle x_{\alpha(i)} | X_{k_i} | x_i \rangle &= \sum_{\substack{\{x_i\} \\ i=1 \dots r}} \dots \langle x_{\alpha^2(1)} | X_{k_{\alpha(1)}} | x_{\alpha(1)} \rangle \langle x_{\alpha(1)} | X_{k_1} | x_1 \rangle \\ &= \prod_{\gamma \in \alpha} \mathrm{Tr} \prod_{i \in \gamma^{-1}} X_{k_i} \end{aligned}$$

where $\gamma \in \alpha$ are the sub-cycles of α and $\prod_{i \in \gamma^{-1}}$ is a product over the numbers in γ^{-1} .

Now consider any set of operators $Y_j \geq 0$, using the spectral decomposition

$Y_j = \sum_k \lambda_{k,j} |k\rangle_j \langle k|_j$ we have

$$\begin{aligned} \left| \text{Tr} \prod_{j=1}^m Y_j \right| &= \left| \sum_{\substack{\{k_j\} \\ j=1\dots m}} \lambda_{k_1,1} \dots \lambda_{k_m,m} \text{Tr} (|k_1\rangle_1 \langle k_1|_1 \dots |k_m\rangle_m \langle k_m|_m) \right| \quad (20) \\ &\leq \sum_{\substack{\{k_j\} \\ j=1\dots m}} \lambda_{k_1,1} \dots \lambda_{k_m,m} = \prod_{j=1}^m \text{Tr} Y_j. \end{aligned}$$

Equality holds only in the following cases:

- If $m = 1$.
- If any of the Y_j 's equals zero.
- If all the Y_j are a multiple of a one-dimensional projection.

To see that there are no other possibilities consider the case where $m \geq 2$ and all Y_j are rank one but they don't have the same eigenvectors. Now the sum in (20) contains the overlap of the eigenvectors which is smaller than one in absolute value because some eigenvectors are not the same. Therefore there is no equality. Finally, consider the case where $m \geq 2$, where all the Y_j have at least rank one and where there exists a j_0 such that Y_{j_0} has rank two or higher. On the RHS of (20), choose the k_j such that all λ_{k_j} are non-zero. For k_{j_0} there are two or more choices and for one of those choices the trace term has to be smaller than one in absolute value. Therefore equality cannot hold in this case.

Returning to the X_j we see that if any $X_j = 0$ we can drop it from our POVM without changing the channel \mathcal{A} . Also if say $X_1 = qX_2$ are multiples of each other then we can combine them to $\tilde{X}_1 = X_1 + X_2$ and $\tilde{\sigma}_1 = (\sigma_1 + q\sigma_2)/(1+q)$ again without changing \mathcal{A} . Therefore we can assume no X_j equals zero and no two X_j are multiples of each other, then equality is only possible if $m = 1$. It follows that

$$\left| \prod_{\gamma \in \alpha} \text{Tr} \prod_{i \in \gamma^{-1}} X_{k_i} \right| \leq \prod_{i=1}^r \text{Tr} X_{k_i},$$

can only be equality if all cycles in α have length one, i.e. $\alpha = \text{id}$.

Combining the last inequality with (19) we get

$$|\mathcal{Q}(\alpha)| \leq \sum_{\substack{\{k_i\} \\ i=1\dots r}} \text{Tr} \left(\prod_{i=1}^r \sigma_{k_i} \right) \prod_{i=1}^r \text{Tr} X_{k_i} = \mathcal{Q}(\text{id})$$

with equality if and only if $\alpha = \text{id}$. And finally,

$$\begin{aligned}
\mathcal{Q}(\text{id}) &= \sum_{\substack{\{k_i\} \\ i=1 \dots r}} \text{Tr} \left(\prod_{i=1}^r \sigma_{k_i} \right) \prod_{i=1}^r \text{Tr} X_{k_i} \\
&= \sum_{\substack{\{k_i\} \\ i=1 \dots r}} \text{Tr} \left(\prod_{i=1}^r \sigma_{k_i} \text{Tr} X_{k_i} \right) \\
&= \text{Tr} \left(\prod_{i=1}^r \sum_k \sigma_k \text{Tr} X_k \right) \\
&= \text{Tr} (\mathcal{A}(\mathbb{1})^r) .
\end{aligned}$$

□

2.6 The qubit depolarizing channel

2.6.1 Evaluating $\mathcal{Q}_{\Delta_\lambda(\alpha)}$

In this section we calculate $\mathcal{Q}_{\Delta_\lambda}(\alpha)$ for $\alpha = \text{id}$ and $\alpha = (1 \dots r)$. As we prove in Lemma 13 in 3.1 one of these two terms is always maximal, $\mathcal{Q}_{\max} = \max \{ \mathcal{Q}_{\Delta_\lambda}(\text{id}), \mathcal{Q}_{\Delta_\lambda}((1 \dots r)) \}$, so they are of particular interest.

We have

$$\mathcal{Q}_{\Delta_\lambda}(\text{id}) = \text{Tr} \Delta_\lambda(\mathbb{1})^r = \text{Tr} \mathbb{1} = 2$$

To evaluate the second \mathcal{Q} -term we consider a slightly more general channel \mathcal{A} with Choi-Jamiolkowski representation

$$\text{Choi}(\Delta_\lambda) = \begin{pmatrix} \mu & 0 & 0 & \lambda \\ 0 & \nu & \kappa & 0 \\ 0 & \kappa & \nu & 0 \\ \lambda & 0 & 0 & \mu \end{pmatrix} .$$

This matrix has diagonal block form with blocks

$$\begin{pmatrix} \mu & \lambda \\ \lambda & \mu \end{pmatrix}, \begin{pmatrix} \nu & \kappa \\ \kappa & \nu \end{pmatrix} .$$

We need to raise the matrix to the r -th power, which gives

$$\begin{pmatrix} \mu & \lambda \\ \lambda & \mu \end{pmatrix}^r = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} (\mu + \lambda)^r & 0 \\ 0 & (\mu - \lambda)^r \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

and similarly for the second matrix. Therefore we get

$$\mathcal{Q}_{\mathcal{A}}((1 \dots r)) = \text{Tr} (\text{Choi}_{\Delta_\lambda}^r) = (\mu + \lambda)^r + (\mu - \lambda)^r + (\nu + \kappa)^r + (\nu - \kappa)^r . \quad (21)$$

For $\mathcal{A} = \Delta_\lambda$ we have $\mu = \frac{1+\lambda}{2}$, $\nu = \frac{1-\lambda}{2}$ and $\kappa = 0$, so

$$\mathcal{Q}_{\Delta_\lambda}((1 \dots r)) = \left(\frac{1+3\lambda}{2} \right)^r + 3 \left(\frac{1-\lambda}{2} \right)^r .$$

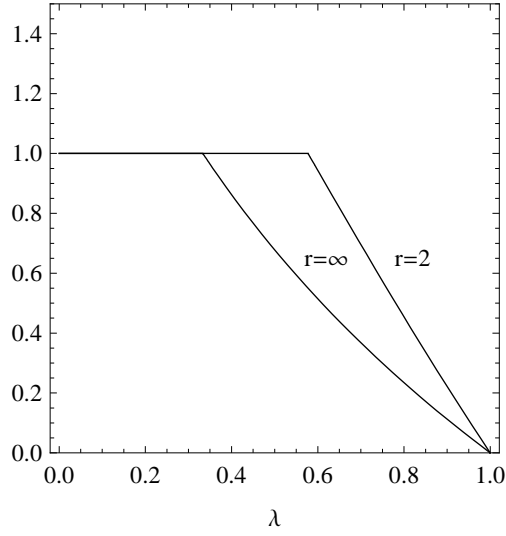


Figure 1: S_2 and S_∞ .

r	c_r	d_r
2	.577	.732
3	.5	.835
4	.458	.878
10	.381	.953
100	.338	.995

Table 1: Range parameters of Theorem 6.

In 4.2 we compute the $\mathcal{Q}((1 \dots r))$ for any dimension $d \geq 2$.

More generally, whenever α is a product of cycles of consecutive numbers the sum factors as shown in Lemma 10, e.g. $\mathcal{Q}((123)(45)) = \frac{1}{d} \mathcal{Q}((123)) \mathcal{Q}((45))$.

2.6.2 Regularized output entropy of $\Delta_\lambda^{\otimes n}$

We are now ready to prove the main theorem.

Theorem 6. (a) For all $r \in \mathbb{N}$, $r \geq 2$, and $\lambda \in [0, 1]$,

$$\beta_r^{\text{reg}}(\Delta_\lambda) = \lim_{n \rightarrow \infty} \frac{1}{n} \beta_r(\Delta_\lambda^{\otimes n}) = \min \left\{ 1, \frac{2r - \log[(1 + 3\lambda)^r + 3(1 - \lambda)^r]}{r - 1} \right\} \quad (22)$$

in particular

$$\beta_2^{\text{reg}}(\Delta_\lambda) = \begin{cases} 1 & \lambda \leq 1/\sqrt{3} \\ 2 - \log(1 + 3\lambda^2) & \lambda > 1/\sqrt{3} \end{cases},$$

$$\beta_\infty^{\text{reg}}(\Delta_\lambda) = \begin{cases} 1 & \lambda \leq 1/3 \\ 2 - \log(1 + 3\lambda) & \lambda > 1/3 \end{cases},$$

where $\beta_\infty^{\text{reg}}(\Delta_\lambda) = \lim_{r \rightarrow \infty} \beta_r^{\text{reg}}(\Delta_\lambda)$.

(b) For all $r \in \mathbb{N}$, $r \geq 2$, and $\lambda \in J_r$,

$$\overline{S}_r^{\text{reg}}(\Delta_\lambda) = \beta_r^{\text{reg}}(\Delta_\lambda) \quad (23)$$

where $J_r = [0, c_r] \cup [d_r, 1] \subset [0, 1]$ for some $0 < c_r < d_r < 1$ (see Table 1).

The regularized output entropy for $r = 2$ and the lower bound of the same for $r = \infty$ are plotted in Figure 1. We expect that (23) holds for all λ . The regularized output entropy is maximal when $2 \geq \left(\frac{1+3\lambda}{2}\right)^r + 3\left(\frac{1-\lambda}{2}\right)^r$ which holds for all r when $\lambda \leq 1/3$. It's interesting to note that $\lambda \leq 1/3$ is also the condition for Δ_λ to be entanglement breaking.

Proof. (a) From 4.2 we know that $\mathcal{Q}(\text{id}) = 2$ and $\mathcal{Q}((1 \dots r)) = \left(\frac{1+3\lambda}{2}\right)^r + 3\left(\frac{1-\lambda}{2}\right)^r$. In Lemma 13 in 3.1 we prove that one of these two \mathcal{Q} -terms yields the maximal value, i.e. for any $\lambda \in [0, 1]$

$$\mathcal{Q}_{\max} = \max \left\{ 2, \left(\frac{1+3\lambda}{2}\right)^r + 3\left(\frac{1-\lambda}{2}\right)^r \right\}.$$

Additionally, Δ_λ is entrywise positive for $\lambda \in [0, 1]$. Therefore we can apply parts (a) and (c) of Theorem 4 and (22) follows immediately.

In the case $r = 2$ we have

$$\frac{2 \cdot 2 - \log[(1+3\lambda)^2 + 3(1-\lambda)^2]}{2-1} = 4 - \log(4 + 12\lambda^2)$$

$$= 2 - \log(1 + 3\lambda^2),$$

and in the case $r = \infty$ we have

$$\begin{aligned} \lim_{r \rightarrow \infty} \beta_r^{\text{reg}}(\Delta_\lambda) &= \lim_{r \rightarrow \infty} \frac{2r - \log[(1+3\lambda)^r + 3(1-\lambda)^r]}{r-1} \\ &= \lim_{r \rightarrow \infty} \frac{2r - r \log(1+3\lambda)}{r-1} \\ &= 2 - \log(1+3\lambda). \end{aligned}$$

(b) Define the functions

$$f(|\phi\rangle) = \text{Tr}(\Delta_\lambda^{\otimes n}(|\phi\rangle\langle\phi|)^r)$$

$$g(x) = \frac{1}{n(1-r)} \log x,$$

and the two series (notice the n dependency in the definition of f)

$$\begin{aligned} a_n &= g(\mathbb{E}f), \\ b_n &= \mathbb{E}g \circ f. \end{aligned}$$

The only difference between a_n and b_n is the position at which the averaging over pure inputs $|\phi\rangle$ takes places. As a result, $a_n = \frac{1}{n}\beta_r(\Delta_\lambda^{\otimes n})$ contains the average moments that have been considered in previous sections, and $b_n = \frac{1}{n}S_r(\Delta_\lambda^{\otimes n})$ is the Renyi output entropy per systems. By Jensen's inequality and because the maximal output entropy is $\log 2 = 1$ we have the bounds $a_n \leq b_n \leq 1$ for any λ . Our goal is to prove that the two series have the same limit for $\lambda \in [0, c_r] \cup [d_r, 1]$.

From part (a) we know

$$\lim_{n \rightarrow \infty} a_n = \frac{r - \log Q_{\max}}{r - 1}.$$

First, choose $c(r)$ such that $2 \geq \left(\frac{1+3\lambda}{2}\right)^r + 3\left(\frac{1-\lambda}{2}\right)^r$ for all $\lambda \in [0, c]$. Then we have $Q_{\max} = 2$, and so

$$\lim_{n \rightarrow \infty} a_n = \lim_{n \rightarrow \infty} b_n = \overline{S}_r^{\text{reg}}(\Delta_\lambda) = 1.$$

Now only (23) remains to be proved for the range $\lambda \in [d, 1]$ with d still to be determined.

In Proposition 18 in 3.2 we prove that the Lipschitz constant η of the map f is bounded by $\eta \leq \sqrt{2}r\kappa^n$ for $\kappa = \lambda + \frac{1-\lambda}{\sqrt{2}}$. From Levy's Lemma (according to Lemma III.1 in [14]) we know that the values of f concentrate around their average

$$\Pr(|f(|\phi\rangle) - \mathbb{E}f| > \alpha_n) \leq 4 \exp\left(-C(k+1)\frac{\alpha_n^2}{\eta^2}\right) =: \epsilon_n \quad (24)$$

with $k = 2 \cdot 2^n - 1$ the (real) dimension of the sphere of input states, $C = (9\pi^3 \ln 2)^{-1}$, and we choose the deviation

$$\alpha_n = \frac{1}{2}N \left(\frac{Q_{\max}}{2^r}\right)^n \approx \frac{1}{2}\mathbb{E}f$$

where $N \geq 1$ is the multiplicity of the maximum of $Q(\alpha)$, that is, α_n is half of the dominant term of $\mathbb{E}f$ (see prove of Theorem 4). Now $\alpha_n \rightarrow 0$ (apart from the special case $\lambda = 1$) and $\mathbb{E}f - \alpha_n > 0$ which is required in a later step. To ensure concentration for large n the exponent $(k+1)\frac{\alpha_n^2}{\eta^2}$ needs to become large. Taking the $2n$ -th root we get

$$\begin{aligned} \left((k+1)\frac{\alpha_n^2}{\eta^2}\right)^{1/2n} &= \left(2 \cdot \sqrt{2}^n \frac{N}{4} \left(\frac{Q_{\max}}{2^r}\right)^{2n} \frac{1}{2r^2 \kappa^{2n}}\right)^{1/2n} \\ &\approx \sqrt{2} \frac{Q_{\max}}{2^r \kappa} \end{aligned}$$

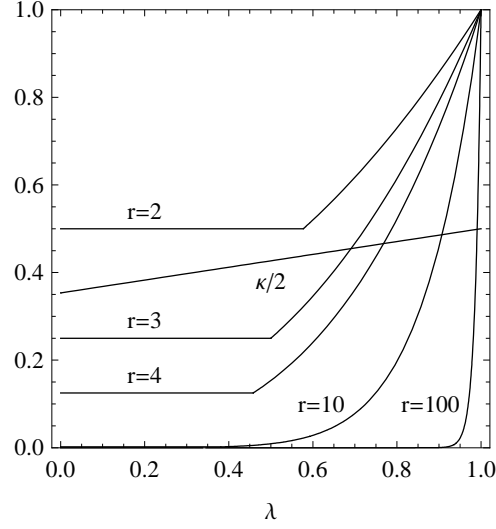


Figure 2: Plots of both sides of (25) for some r values to determine the range of validity.

where the approximation becomes equality in the large n limit. If this term is larger than 1 we have the required divergence, this gives the following inequality

$$\frac{Q_{\max}}{2^r} > \frac{\kappa}{\sqrt{2}}. \quad (25)$$

This condition gives the lower bound d_r for the λ values for which (23) holds. Both sides of the inequality are plotted in Figure 2. For some values of r the range parameters can be found in Table 1.

To transform (24) into a statement about the range of g – and whence about a_n – we need a condition that implies $|f(|\phi\rangle) - \mathbb{E}f| > \alpha_n$. We set $\alpha'_n = \frac{1}{n(r-1)} \frac{\alpha_n}{\mathbb{E}f - \alpha_n} > |g(\mathbb{E}f) - g(\mathbb{E}f - \alpha_n)| > |g(\mathbb{E}f) - g(\mathbb{E}f + \alpha_n)|$ where the second inequality follows from convexity of g . Then for any x we have

$$|g(\mathbb{E}f) - g(x)| > \alpha'_n \Rightarrow |\mathbb{E}f - x| > \alpha_n.$$

Therefore, the values of $g \circ f$ concentrate around $a_n = g(\mathbb{E}f)$

$$\Pr(|g \circ f(|\phi\rangle) - a_n| > \alpha'_n) \leq \epsilon_n.$$

For α'_n we calculate

$$\begin{aligned} \lim_{n \rightarrow \infty} \alpha'_n &= \lim_{n \rightarrow \infty} \frac{1}{n(r-1)} \frac{\alpha_n}{\mathbb{E}f - \alpha_n} \\ &= \lim_{n \rightarrow \infty} \frac{1}{n(r-1)} \frac{1/2}{1/2} \\ &= 0. \end{aligned}$$

To find an upper bound on b_n assume all $|\phi\rangle \in (g \circ f)^{-1}[a_n - \alpha'_n, a_n + \alpha'_n] = B$ map to the maximal value $a_n + \alpha'_n$ and all $|\phi\rangle \in B^c$ map to the maximal value 1. This gives a possible range for the average

$$b_n \in [a_n, (1 - \epsilon_n)(a_n + \alpha'_n) + \epsilon_n \cdot 1]. \quad (26)$$

For $\lambda \in [d, 1]$ both ϵ_n and α'_n tend to 0 for large n and with (26) and because we know the limit of a_n we have

$$\lim_{n \rightarrow \infty} a_n = \lim_{n \rightarrow \infty} b_n = \frac{r - \log \mathcal{Q}_{\max}}{r - 1}.$$

□

2.6.3 Output entropy of random sequences

Considering sequences of random pure input states $|\phi_n\rangle$ with increasing dimension 2^n we have the following statement.

Proposition 8. *Let $|\phi_n\rangle \in \mathbb{C}^{2^n}$ be a sequence of random pure states and*

$$c_n = \frac{1}{n} S_r(\Delta_\lambda^{\otimes n}(|\phi_n\rangle\langle\phi_n|))$$

the sequence of output Renyi entropies per system. Then

$$c_n \xrightarrow{\text{a.s.}} \frac{r - \log \mathcal{Q}_{\max}}{r - 1}$$

if λ is restricted as in Theorem 6.

Proof. Define $f_n(|\phi\rangle) = \frac{1}{n} S_r(\Delta_\lambda^{\otimes n}(|\phi\rangle\langle\phi|))$, the limit value $c = \frac{r - \log \mathcal{Q}_{\max}}{r - 1} = \lim_{n \rightarrow \infty} \mathbb{E} f_n$, and $\delta_n = |\mathbb{E} f_n - c|$. Then for any $\epsilon > 0$

$$\begin{aligned} \Pr(|f_n(|\phi\rangle) - c| > \epsilon) &\leq \Pr(|f_n(|\phi\rangle) - \mathbb{E} f_n| > \epsilon - \delta_n) \\ &\leq 4 \exp\left(-C(k+1) \frac{(\epsilon - \delta_n)^2}{\eta^2}\right), \end{aligned} \quad (27)$$

with C , $k = 2 \cdot 2^n - 1$, and $\eta \leq \sqrt{2} r \kappa^n$ with $0 < \kappa < 1$ as in the proof of Theorem 6. If $\lambda \in J_r$ then $\lim \delta_n = 0$. When n becomes large then k becomes large, η becomes small, and $\epsilon - \delta_n$ is close to $\epsilon > 0$. Therefore the probabilities in (27) become small. Let N be such that for $n > N$ we have $\delta_n < \epsilon/2$. Then we have the bound

$$\begin{aligned} \sum_{n=1}^{\infty} \Pr(|f_n(|\phi_n\rangle) - c| > \epsilon) &\leq \sum_{n=1}^{\infty} 4 \exp\left(-C(k+1) \frac{(\epsilon - \delta_n)^2}{\eta^2}\right) \\ &\leq N + \sum_{n>N} 4 \exp\left(-C(k+1) \frac{(\epsilon/2)^2}{\eta^2}\right) \\ &\leq N + \sum_{n>N} 4 \exp\left(-\tilde{C}(2/\kappa^2)^n\right) \\ &< \infty, \end{aligned}$$

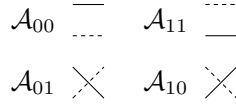


Figure 3: The diagrams for the \mathcal{A}_{xy} matrices.

$$\mathcal{A}_{00}\mathcal{A}_{01}\mathcal{A}_{10}\mathcal{A}_{11}\mathcal{A}_{10}$$



Figure 4: A product of \mathcal{A}_{xy} matrices and the corresponding diagram.

with $\tilde{C} = C \cdot 2 \cdot \frac{(\epsilon/2)^2}{(\sqrt{2}r)^2}$ independent of n . The Proposition follows by the lemma of Borel-Cantelli. \square

3 Proof of lemmas

3.1 Maximal \mathcal{Q} for Δ_λ

Before proving Proposition 13 which is required for the proof of Theorem 6 we introduce some new notation and we present four lemmas. We only prove our Lemmas in two dimensions, but similar results will hold for higher dimensions.

For the most part we consider a channel \mathcal{A} slightly more general than the depolarization channel with

$$\begin{aligned} \mathcal{A}_{00} &= \begin{pmatrix} \mu & 0 \\ 0 & \nu \end{pmatrix} & \mathcal{A}_{11} &= \begin{pmatrix} \nu & 0 \\ 0 & \mu \end{pmatrix} \\ \mathcal{A}_{10} &= \begin{pmatrix} 0 & \kappa \\ \lambda & 0 \end{pmatrix} & \mathcal{A}_{01} &= \begin{pmatrix} 0 & \lambda \\ \kappa & 0 \end{pmatrix} \end{aligned},$$

where $\kappa, \lambda, \mu, \nu \in \mathbb{R}_0^+$ and $\mu \geq \nu, \lambda \geq \kappa$. We call this channel the *two-rail channel*. Remember that in the case $\mathcal{A} = \Delta_\lambda$ we have $\mu = \frac{1+\lambda}{2}, \nu = \frac{1-\lambda}{2}$ and $\kappa = 0$. Because all these matrix entries are positive, so are the \mathcal{Q} -terms, and therefore the largest positive \mathcal{Q} -term will yield \mathcal{Q}_{\max} .

We think of these matrices as the diagrams in Figure 3. We refer to the lines as rails and to their vertical position as their track (starting with track 0)

A product looks like the diagram in Figure 4, notice that we read from right to left, the same way that matrix multiplication applies. Consider a vector multiplying with this product, the diagram can be thought of as presenting two rails along which the two entries of the vector pass through to the left. On the way, continuous lines multiply the entries with factors μ or λ , dashed lines with a factor ν or κ . Because the products are inside a trace, they will only contribute, if the rail starting at the top on the right, ends on the top at the left (giving the

$\langle 0 | \dots | 0 \rangle$ contribution), and the same for the rail starting at the bottom (giving the $\langle 1 | \dots | 1 \rangle$ contribution). Using this, we can compare contributions to $\mathcal{Q}(\alpha)$ for different α .

We rewrite

$$\mathcal{Q}(\alpha) = \sum_{\delta \in \mathcal{D}(\alpha)} \text{Tr}(\delta)$$

where $\mathcal{D}(\alpha)$ is the set of 2^r diagrams corresponding to α and we identify the diagram δ with the corresponding matrix.

Furthermore, for a diagram δ we define δ^1 to be the unchanged diagram and δ^{-1} to be the horizontally reflected diagram. If we take $\mathcal{D}_0(\alpha)$ to be the set of all diagrams in $\mathcal{D}(\alpha)$ that start high (at track 0) then obviously $\mathcal{D}(\alpha) = \mathcal{D}_0(\alpha) \cup \mathcal{D}_0^{-1}(\alpha)$ provides a convenient splitting of the sum in $\mathcal{Q}(\alpha)$.

We say an α is *non-overlapping* if all its cycles permute consecutive numbers, e.g. $\alpha = (123)(45)(678)$. For such α we write $\alpha = \alpha_1 \dots \alpha_s$, where all the α_i are the cycles. We define a product for diagrams by simply concatenating them. With this product we get $\mathcal{D}(\alpha) = \mathcal{D}(\alpha_1) \dots \mathcal{D}(\alpha_s)$.

Lemma 9. *Consider a diagram $\delta \in \mathcal{D}((1 \dots r))$ of a two-rail channel. Then we have*

$$\delta + \delta^{-1} \propto \mathbb{1}.$$

Proof. In any diagram $\delta \in \mathcal{D}((1 \dots r))$ one rail starts and ends at track 0 and the other starts and ends at track 1. That means the corresponding matrix is diagonal. Reflecting the diagram simply means exchanging the two diagonal entries, and if we sum $\delta + \delta^{-1}$ then the diagonal entries both have the same sum, i.e. it is proportional to unity. \square

Lemma 10. *Let $\alpha = \alpha_1 \dots \alpha_s \in \text{Sym}(r)$ be a non-overlapping permutation consisting of s cycles. Then the $\mathcal{Q}(\alpha)$ of a two-rail channel factors like*

$$\mathcal{Q}(\alpha) = \frac{1}{2^{s-1}} \mathcal{Q}(\alpha_1) \dots \mathcal{Q}(\alpha_s).$$

Proof. First we use the fact, that the diagrams for a non-overlapping α can be split between cycles, i.e. the sum splits into sums over separate diagrams,

$$\begin{aligned} \mathcal{Q}(\alpha) &= \sum_{\delta \in \mathcal{D}(\alpha)} \text{Tr}(\delta) \\ &= \sum_{\delta_1 \in \mathcal{D}(\alpha_1)} \dots \sum_{\delta_s \in \mathcal{D}(\alpha_s)} \text{Tr}(\delta_1 \dots \delta_s) \\ &= \text{Tr} \left(\sum_{\delta_1 \in \mathcal{D}(\alpha_1)} \delta_1 \dots \sum_{\delta_s \in \mathcal{D}(\alpha_s)} \delta_s \right). \end{aligned}$$

Using the splitting $\mathcal{D}(\alpha) = \mathcal{D}_0(\alpha) \cup \mathcal{D}_0^{-1}(\alpha)$ and Lemma 9 we have

$$\sum_{\delta \in \mathcal{D}(\alpha_i)} \delta = \sum_{\delta \in \mathcal{D}_0(\alpha_i)} \delta + \delta^{-1} \propto \mathbb{1}$$

for all s cycles α_i . Therefore, we can split the single trace into s separate traces

$$\begin{aligned}\mathcal{Q}(\alpha) &= \frac{1}{2^{s-1}} \sum_{\delta_1 \in \mathcal{D}(\alpha_1)} \text{Tr}(\delta_1) \cdots \sum_{\delta_s \in \mathcal{D}(\alpha_s)} \text{Tr}(\delta_s) \\ &= \frac{1}{2^{s-1}} \mathcal{Q}(\alpha_1) \cdots \mathcal{Q}(\alpha_s).\end{aligned}$$

□

Lemma 11. *For $\mathcal{A} = \Delta_\lambda$ the depolarizing channel in two dimensions, \mathcal{Q} restricted to non-overlapping permutations $\alpha \in \text{Sym}(r)$ is either maximal when $\alpha = (1 \dots r)$ or when $\alpha = \text{id}$.*

Proof. Remember from 2.6.1 that $\mathcal{Q}((1 \dots r)) = (\mu + \lambda)^r + 3\nu^r$. It is convenient to introduce

$$f(x) = (\mu + \lambda)^x + 3\nu^x,$$

as a function with range \mathbb{R} . For a pure $\alpha = \alpha_1 \dots \alpha_s$ according to Lemma 10 we get

$$\mathcal{Q}(\alpha) = \frac{1}{2^{s-1}} \prod_{i=1}^s f(|\alpha_i|). \quad (28)$$

with $|\alpha_i|$ denoting the length of a cycle.

In the following we keep the number s of cycles and the total length r of the permutation invariant. Now, if we increase the length of one cycle and decrease the length of another, we are only changing two factors in the product, $f(x)f(k-x)$, where x is the length of the first cycle and k the (invariant) sum of the lengths of both cycles. We rewrite

$$\begin{aligned}f(x)f(k-x) &= ((\mu + \lambda)^x + 3\nu^x)((\mu + \lambda)^{k-x} + 3\nu^{k-x}) \\ &= (\mu + \lambda)^k + 9\nu^k + 3\nu^k \left(\frac{\mu + \lambda}{\nu} \right)^x + 3(\mu + \lambda)^k \left(\frac{\nu}{\mu + \lambda} \right)^x.\end{aligned}$$

Because of $\mu + \lambda, \nu \geq 0$ the function $f(x)f(k-x)$ is convex in x . Therefore it is maximal at the boundaries, i.e. when one cycle has the minimal length of 1. If we repeat this procedure $s-1$ times we end up with one large cycle of length $t = r - s + 1$ while all other cycles are of length 1. In every step we increase \mathcal{Q} , so we get the bound

$$\mathcal{Q}(\alpha) \leq \mathcal{Q}((1)(2) \dots (s-1)(s \dots r)) = \frac{1}{2^{s-1}} f(1)^{s-1} f(t) = f(t) \quad (29)$$

for non-overlapping permutations α consisting of s cycles.

Now compare the upper bounds given by (29) for permutations of the same total length r but different number of cycles s . This is the same as varying t . Because f is convex we get a maximal upper bound if t is minimal or maximal. The minimal value $t = 1$ is achieved when $s = r$ and all the cycles are of length 1. Then (29) becomes an equality - there are no steps necessary

in the maximization procedure - and we have $\mathcal{Q}(\text{id}) = f(1) = 2$. The maximal value $t = r$ is achieved when α is simply one large cycle. Again (29) becomes equality, and $\mathcal{Q}(\alpha) = f(r)$. One of these upper bounds is the highest upper bound possible in (29), and because they are achieved by $\mathcal{Q}(\text{id})$ and $\mathcal{Q}((1 \dots r))$ we know that one of these $\mathcal{Q}(\alpha)$ is maximal over $\alpha \in \text{Sym}(r)$. \square

Lemma 12. *Let \mathcal{A} be a two-rail channel, and $[\beta]$ the conjugacy class of a permutation. Then \mathcal{Q} restricted to $[\beta]$ is maximal on non-overlapping members $\alpha \in [\beta]$.*

Proof. Let $\alpha = \alpha_1 \dots \alpha_s$ be a non-overlapping member of the class and $\beta = \gamma \alpha \gamma^{-1}$ be any other member of the class. First, remember

$$\mathcal{Q}(\alpha) = \sum_{\substack{\{x_i\} \\ i=1 \dots r}} \text{Tr} \left(\prod_{i=1}^r \mathcal{A}_{x_i x_{\alpha(i)}} \right) = \sum_{\delta_1 \in \mathcal{D}(\alpha_1)} \dots \sum_{\delta_s \in \mathcal{D}(\alpha_s)} \text{Tr}(\delta_1 \dots \delta_s). \quad (30)$$

With the first way of writing $\mathcal{Q}(\alpha)$ in mind we define a 1-1-mapping between terms in the sum of $\mathcal{Q}(\alpha)$ and $\mathcal{Q}(\beta)$ via a mapping of indices $x_i \rightarrow x_{\gamma^{-1}(i)}$ (or $x_{\gamma(i)} \rightarrow x_i$). Then the products of matrices are mapped like

$$\begin{aligned} \mathcal{A}_{x_1 x_{\alpha(1)}} \dots \mathcal{A}_{x_r x_{\alpha(r)}} &\rightarrow \mathcal{A}_{x_{\gamma(1)} x_{\gamma(\beta(1))}} \dots \mathcal{A}_{x_{\gamma(r)} x_{\gamma(\beta(r))}} \\ &= \mathcal{A}_{x_{\gamma(1)} x_{\alpha(\gamma(1))}} \dots \mathcal{A}_{x_{\gamma(r)} x_{\alpha(\gamma(r))}}. \end{aligned}$$

In terms of diagrams this corresponds to permuting the “tiles” (crossings or straight pieces) $\mathcal{A}_{x_i x_{\alpha(i)}}$ according to the permutation γ . Some examples are shown in Appendix A. In particular, this mapping does not change the number of any kind of tile \mathcal{A}_{00} , \mathcal{A}_{01} , \mathcal{A}_{10} or \mathcal{A}_{11} .

Now, consider the second way of writing $\mathcal{Q}(\alpha)$ in (30) and split the sums over subdiagrams $\mathcal{D}(\alpha_i)$ according to $\mathcal{D}(\alpha_i) = \mathcal{D}_0(\alpha_i) \cup \mathcal{D}_0^{-1}(\alpha_i)$

$$\mathcal{Q}(\alpha) = \sum_{\delta_1 \in \mathcal{D}_0(\alpha_1)} \dots \sum_{\delta_s \in \mathcal{D}_0(\alpha_s)} \sum_{\substack{\{t_i = \pm 1\} \\ i=1 \dots s}} \text{Tr}(\delta_1^{t_1} \dots \delta_s^{t_s}).$$

We consider a subsum

$$\sum_{\substack{\{t_i = \pm 1\} \\ i=1 \dots s}} \text{Tr}(\delta_1^{t_1} \dots \delta_s^{t_s})$$

for fixed subdiagrams $\delta_i \in \mathcal{D}_0(\alpha_i)$. In the following we will prove that the contribution of this subsum to $\mathcal{Q}(\alpha)$ is larger or equal to the contribution of the subsum of the corresponding diagrams to $\mathcal{Q}(\beta)$. From this it immediately follows that $\mathcal{Q}(\alpha) \geq \mathcal{Q}(\beta)$.

It follows a general proof of the inequality between corresponding subsums. For illustration one subsum is evaluated in full detail with diagrams in Appendix A.

For the non-overlapping permutation α in the subdiagrams δ_i all the straight lines and dashed lines are aligned, i.e. we have a weak and a strong rail. Let m_i be the number of crossings and n_i the number of straight pieces in δ_i . The strong rail in subdiagram δ_i contributes a factor $\lambda^{m_i} \mu^{n_i}$ and the weak rail contributes the factor $\kappa^{m_i} \nu^{n_i}$. Summing over reflections the subsum equals

$$2 \prod_{i=1}^s \lambda^{m_i} \mu^{n_i} + \kappa^{m_i} \nu^{n_i}.$$

On the other hand, for the possibly overlapping β some of the tiles are permuted and for one particular subdiagram, not all the strong rail pieces might be on the same rail. Let \tilde{m}_i be the number of crossings that are thus misaligned and \tilde{n}_i the number of straight pieces that are misaligned. The two rails in subdiagram δ_i now contribute the factors $\lambda^{m_i - \tilde{m}_i} \kappa^{\tilde{m}_i} \mu^{n_i - \tilde{n}_i} \nu^{\tilde{n}_i}$ and $\lambda^{\tilde{m}_i} \kappa^{m_i - \tilde{m}_i} \mu^{\tilde{n}_i} \nu^{n_i - \tilde{n}_i}$. Summing over reflections the subsum adding to $\mathcal{Q}(\beta)$ equals

$$2 \prod_{i=1}^s \lambda^{m_i - \tilde{m}_i} \kappa^{\tilde{m}_i} \mu^{n_i - \tilde{n}_i} \nu^{\tilde{n}_i} + \lambda^{\tilde{m}_i} \kappa^{m_i - \tilde{m}_i} \mu^{\tilde{n}_i} \nu^{n_i - \tilde{n}_i}.$$

Because $\lambda \geq \kappa$ and $\mu \geq \nu$ it follows that $\lambda^{m_i} \mu^{n_i} + \kappa^{m_i} \nu^{n_i} \geq \lambda^{m_i - \tilde{m}_i} \kappa^{\tilde{m}_i} \mu^{n_i - \tilde{n}_i} \nu^{\tilde{n}_i} + \lambda^{\tilde{m}_i} \kappa^{m_i - \tilde{m}_i} \mu^{\tilde{n}_i} \nu^{n_i - \tilde{n}_i}$ (the strong and weak rail dominate the two mixed rails), and hence we have the desired inequality between corresponding contributions to $\mathcal{Q}(\alpha)$ and $\mathcal{Q}(\beta)$. \square

Proposition 13. *For $\mathcal{A} = \Delta_\lambda$ the depolarizing channel in two dimensions, $\mathcal{Q}(\alpha)$ is either maximal when $\alpha = (1 \dots r)$ or when $\alpha = \text{id}$.*

Proof. First, consider permutations that consist of non-overlapping α , Lemma 11 proves that either $\alpha = \text{id}$ or $\alpha = (1 \dots r)$ yields the maximum $\mathcal{Q}(\alpha)$ amongst these permutations. Every conjugacy class has a non-overlapping representant, and Lemma 12 states that these have maximal \mathcal{Q} -value. Therefore either $\alpha = \text{id}$ or $\alpha = (1 \dots r)$ yield the maximal \mathcal{Q} -value amongst all the permutations $\alpha \in \text{Sym}(r)$. \square

3.2 Bound on Lipschitz constant

We derive an upper bound on the Lipschitz constant of the function $f : S^{2^{n+1}-1} \rightarrow \mathbb{R}$

$$f(|\phi\rangle) = \text{Tr}(\Delta_\lambda^{\otimes n}(|\phi\rangle\langle\phi|)^r)$$

with respect to the Euclidean norm on $S^{2^{n+1}-1} \subset \mathbb{R}^{2^{n+1}}$.

We divide the function into four steps and prove bounds on the Lipschitz

constants for each step. The splitting is $f = d \circ c \circ b \circ a$ with

$$\begin{aligned} a : |\phi\rangle &\rightarrow |\phi\rangle\langle\phi| = \rho \\ b : \rho &\rightarrow \Delta_\lambda^{\otimes n}(\rho) = \rho' \\ c : \rho' &\rightarrow \text{eigenvalues of } \rho' = \vec{v} \\ d : \vec{v} &\rightarrow \sum_i v_i^r. \end{aligned}$$

Let \mathcal{M}_m be the space of complex $m \times m$ matrices containing the set of states $\mathcal{S}_m = \{\rho \in \mathcal{M}_m : \rho = \rho^*, \text{Tr } \rho = 1\}$.

Lemma 14. *Let $a(|\phi\rangle) = |\phi\rangle\langle\phi|$ a map $\mathbb{C}^m \rightarrow \mathcal{M}_m$ where $\langle\phi|\phi\rangle = 1$. Then the Lipschitz constant of a with respect to the Euclidean norm in the domain and the Frobenius norm $\|M\|_2 = \text{Tr}(MM^*)^{1/2}$ in the range is upper bounded by $\sqrt{2}$.*

Proof. For $\langle\phi|\phi\rangle = \langle\psi|\psi\rangle = 1$ set $c = \langle\phi|\psi\rangle$. Now

$$\| |\phi\rangle\langle\phi| - |\psi\rangle\langle\psi| \|_2^2 = \langle\phi|\phi\rangle - \langle\phi|\psi\rangle - \langle\psi|\phi\rangle + \langle\psi|\psi\rangle = 2 - 2\Re(c),$$

and

$$\| |\phi\rangle\langle\phi| - |\psi\rangle\langle\psi| \|_2^2 = \langle\phi|\phi\rangle^2 - 2|\langle\phi|\psi\rangle|^2 + \langle\psi|\psi\rangle^2 = 2 - 2|c|^2.$$

Because of $\Re(c) \leq |c|$ and the inequality derived as follows

$$\begin{aligned} (1 - |c|)^2 &\geq 0 \\ 2 - 2|c| &\geq 1 - |c|^2 \\ 2(2 - 2|c|) &\geq 2 - 2|c|^2, \end{aligned}$$

we arrive at $2(2 - 2\Re(c)) \geq 2 - 2|c|^2$ or $\sqrt{2}\| |\phi\rangle\langle\phi| - |\psi\rangle\langle\psi| \|_2 \geq \| |\phi\rangle\langle\phi| - |\psi\rangle\langle\psi| \|_2$. \square

Lemma 15. *Let $b(\rho) = \Delta_\lambda^{\otimes n}(\rho)$ a map $\mathcal{S}_{2^n} \rightarrow \mathcal{S}_{2^n}$, then the Lipschitz constant of b with respect to the Frobenius norm in domain and range is upper bounded by κ^n , where $\kappa = \lambda + \frac{1-\lambda}{\sqrt{2}}$ so $0 < \kappa < 1$ for $0 < \lambda < 1$.*

Proof. It is useful to use the notation

$$\Delta_\lambda^{\otimes n} = \sum_{J \subset \mathbb{Z}_n} \lambda^{n-|J|} (1-\lambda)^{|J|} \text{Tr}_J \otimes \left(\frac{\mathbb{1}}{2} \right)^{\otimes |J|}$$

where Tr_J is the partial trace over the systems with indices in J . Notice that we use a loose notation of the tensor product as the systems that are partially traced out and replaced by the totally mixed states are not necessarily all on the right side of the tensor product.

We will bound the operator norm $\|\mathcal{A}\|_{op} := \sup_{\rho \in \mathcal{S}} \frac{\|\mathcal{A}(\rho)\|_2}{\|\rho\|_2}$ where the supremum is over the set $\mathcal{S} = \{ |\phi\rangle\langle\phi| - |\psi\rangle\langle\psi| \mid |\phi\rangle, |\psi\rangle \in \mathcal{S}_{2^n-1} \}$. Because of

$$\|\mathcal{A}(\rho) - \mathcal{A}(\tau)\|_2 = \|\mathcal{A}(\rho - \tau)\|_2 \leq \|\mathcal{A}\|_{op} \|\rho - \tau\|_2$$

for a linear map \mathcal{A} bounding $\|\mathcal{A}\|_{op}$ immediately gives a bound of the Lipschitz constant as well. Now

$$\begin{aligned}
\|\Delta_\lambda^{\otimes n}\|_{op} &\leq \sum_J \lambda^{n-|J|} (1-\lambda)^{|J|} \left\| \text{Tr}_J \otimes \left(\frac{\mathbb{1}}{2}\right)^{\otimes |J|} \right\|_{op} \\
&\leq \sum_J \lambda^{n-|J|} \left(\frac{1-\lambda}{\sqrt{2}}\right)^{|J|} \\
&= \sum_{k=0}^n \binom{n}{k} \lambda^{n-k} \left(\frac{1-\lambda}{\sqrt{2}}\right)^k \\
&= \left(\lambda + \frac{1-\lambda}{\sqrt{2}}\right)^n = \kappa^n,
\end{aligned}$$

where we used the bound

$$\begin{aligned}
\left\| \text{Tr}_{\{1\dots k\}} \otimes \left(\frac{\mathbb{1}}{2}\right)^{\otimes k} \right\|_{op}^2 &= \sup_{\rho \in S} \frac{\text{Tr} \left(\text{Tr}_{\{1\dots k\}} \rho \otimes \left(\frac{\mathbb{1}}{2}\right)^{\otimes k} \right)^2}{\text{Tr} \rho^2} \\
&= \text{Tr} \left(\left(\frac{\mathbb{1}}{2}\right)^{\otimes k} \right)^2 \sup_{\rho \in S} \frac{\text{Tr} \left(\text{Tr}_{\{1\dots k\}} \rho \right)^2}{\text{Tr} \rho^2} \quad (31) \\
&= \left(\frac{1}{2}\right)^k.
\end{aligned}$$

The supremum in (31) was evaluated as follows. First, consider that $\rho = |\phi\rangle\langle\phi| - |\psi\rangle\langle\psi|$ can be written as $\rho = \alpha|0\rangle\langle 0| - \alpha|1\rangle\langle 1|$ with $0 \leq \alpha \leq 1$, where $|0\rangle$ and $|1\rangle$ are orthonormal states. Then the supremum runs over all possible orientations of $|0\rangle$ and $|1\rangle$

$$\begin{aligned}
\sup_{\rho \in S} \frac{\text{Tr} \left(\text{Tr}_{\{1\dots k\}} \rho \right)^2}{\text{Tr} \rho^2} &= \sup_{|0\rangle, |1\rangle} \frac{\text{Tr} \left(\text{Tr}_{\{1\dots k\}} (\alpha|0\rangle\langle 0| - \alpha|1\rangle\langle 1|) \right)^2}{\text{Tr} (\alpha|0\rangle\langle 0| - \alpha|1\rangle\langle 1|)^2} \\
&= \sup_{|0\rangle, |1\rangle} \frac{\text{Tr} \left(\text{Tr}_{\{1\dots k\}} (|0\rangle\langle 0| - |1\rangle\langle 1|) \right)^2}{\text{Tr} (|0\rangle\langle 0| - |1\rangle\langle 1|)^2} \\
&= \sup_{|0\rangle, |1\rangle} \frac{\text{Tr} \left(\text{Tr}_{\{1\dots k\}} (|0\rangle\langle 0| - |1\rangle\langle 1|) \right)^2}{2}.
\end{aligned}$$

Now assume $\rho_0 = \text{Tr}_{\{1\dots k\}} |0\rangle\langle 0|$ and $\rho_1 = \text{Tr}_{\{1\dots k\}} |1\rangle\langle 1|$ are arbitrary density matrices. Then

$$\begin{aligned}
\sup_{|0\rangle, |1\rangle} \frac{\text{Tr} \left(\text{Tr}_{\{1\dots k\}} (|0\rangle\langle 0| - |1\rangle\langle 1|) \right)^2}{2} &= \sup_{\rho_0, \rho_1} \frac{\text{Tr} (\rho_0 - \rho_1)^2}{2} \\
&= \sup_{\rho_0, \rho_1} \frac{\text{Tr} (\rho_0^2 + \rho_1^2 - 2\rho_0\rho_1)}{2} = 1.
\end{aligned}$$

The last equality follows from the fact, that $\text{Tr } \rho_{0,1}^2 \leq 1$ and $\text{Tr } \rho_0 \rho_1 \geq 0$. The suprema are achieved when $|\phi\rangle = |00\rangle$ and $|\psi\rangle = |11\rangle$ so that $\rho_0 = |0\rangle\langle 0|$ and $\rho_1 = |1\rangle\langle 1|$. \square

Remark 16. Let $c : \mathcal{S}_m \rightarrow \mathbb{R}^m$ be the map that sends density matrices to their eigenvalues, ordered high to low. The fact that the Lipschitz constant of c is upper bounded by 1 is equivalent to the Hoffman-Wielandt inequality [19]

$$\|\rho - \tau\|_2 \geq \|c(\rho) - c(\tau)\|_2.$$

Lemma 17. *Let $d(\vec{v}) = \sum_i v_i^r$ a map from $\{\vec{w} \in \mathbb{R}_+^m | \sum_i w_i = 1\}$ to \mathbb{R} . Then the Lipschitz constant of d is upper bounded by r .*

Proof. We have $\frac{\partial d}{\partial v_j} = r v_j^{r-1}$ so

$$\sup_{\vec{v} \in \text{Dom } d} |\vec{\nabla} d|^2 = \sup \sum r^2 v_i^{2(r-1)} \leq r^2.$$

By integration we get

$$|d(\vec{v}) - d(\vec{w})| \leq r \|\vec{v} - \vec{w}\|_2.$$

\square

Proposition 18. *The Lipschitz constant η of $\text{Tr}(\Delta_\lambda^{\otimes n}(|\phi\rangle\langle\phi|)^r)$, with respect to the Euclidean norm in the domain, is upper bounded by $\sqrt{2}r\kappa^n$, with κ as in Lemma 15.*

Proof. From $\text{Tr}(\Delta_\lambda^{\otimes n}(|\phi\rangle\langle\phi|)^r) = d \circ c \circ b \circ a(|\phi\rangle)$ with a, b, c and d as defined in Lemmas/Remark 14-17. Thus, we simply combine the upper bounds of the lemmas and get the bound $\sqrt{2} \cdot \kappa^n \cdot 1 \cdot r$. \square

4 Other results

4.1 $\mathcal{Q}((1 \dots r))$ for a more general qubit channel

For a channel \mathcal{A} that maps that scales the Bloch sphere like

$$\vec{n} \rightarrow \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} \vec{n},$$

the Choi-Jamiolkowski representation is

$$\text{Choi}(\Delta_\lambda) = \begin{pmatrix} \mu & 0 & 0 & \lambda \\ 0 & \nu & \kappa & 0 \\ 0 & \kappa & \nu & 0 \\ \lambda & 0 & 0 & \mu \end{pmatrix},$$

i.e. this is a two-rail channel with $\mu = \frac{1+\lambda_3}{2}$, $\nu = \frac{1-\lambda_3}{2}$, $\lambda = \frac{\lambda_1+\lambda_2}{2}$ and $\kappa = \frac{\lambda_1-\lambda_2}{2}$. Therefore we have

$$\mathcal{Q}(\text{id}) = \text{Tr } \mathcal{A}(\mathbb{1})^r = \text{Tr } \mathbb{1} = 2$$

and according to (21)

$$\begin{aligned} \mathcal{Q}((1 \dots r)) &= (\mu + \lambda)^r + (\mu - \lambda)^r + (\nu + \kappa)^r + (\nu - \kappa)^r \\ &= \left(\frac{1 + \lambda_1 + \lambda_2 + \lambda_3}{2} \right)^r + \left(\frac{1 - \lambda_1 - \lambda_2 + \lambda_3}{2} \right)^r \\ &\quad + \left(\frac{1 + \lambda_1 - \lambda_2 - \lambda_3}{2} \right)^r + \left(\frac{1 - \lambda_1 + \lambda_2 - \lambda_3}{2} \right)^r. \end{aligned}$$

Assuming $\mathcal{Q}_{\max} = \max\{2, \mathcal{Q}((1 \dots r))\}$ the output is maximally mixed if $|\lambda_1| + |\lambda_2| + |\lambda_3| \leq 1$. But this is exactly the condition for \mathcal{A} to be entanglement breaking, according to Theorem 3 in [6].

4.2 $\mathcal{Q}((1 \dots r))$ of Δ_λ for any dimension d

For the d -dimensional depolarizing channel we have $\mathcal{Q}(\text{id}) = \text{Tr } \mathbb{1} = d$ as usual. Furthermore, the Choi-Jamiolkowski representation consists of two blocks, one d -dimensional block with diagonal entries $\mu = \frac{1+\lambda}{d}$ and off-diagonal entries λ , and another block that is $\nu = \frac{1-\lambda}{d}$ times identity on the other $d^2 - d$ dimensions. For example for $d = 3$ the representation looks like

$$\text{Choi}(\Delta_\lambda) = \begin{pmatrix} \mu & & & \lambda & & \lambda \\ & \nu & & & & \\ & & \nu & & & \\ & & & \nu & & \\ \lambda & & & \mu & & \lambda \\ & & & & \nu & \\ & & & & & \nu \\ & & & & & & \nu \\ \lambda & & & \lambda & & & \mu \end{pmatrix}.$$

The eigenvalues of the first block are $\mu - \lambda$ with $(d-1)$ -multiplicity and a single eigenvalue $\mu + (d-1)\lambda$. Therefore,

$$\begin{aligned} \mathcal{Q}((1 \dots r)) &= \text{Tr } \text{Choi}(\Delta_\lambda)^r \\ &= (\mu + (d-1)\lambda)^r + (d-1)(\mu - \lambda)^r + (d^2 - d)\nu^r \\ &= \left(\frac{1 + (d^2 - 1)\lambda}{d} \right)^r + (d^2 - 1) \left(\frac{1 - \lambda}{d} \right)^r. \end{aligned}$$

This result agrees with the result for $d = 2$ found in 2.6.1. The critical value is $\lambda = \frac{1}{d+1}$, below this value the average output is maximally mixed.

5 Conclusion

We found a general explicit form for $\beta_r(\mathcal{A})$ depending on the function $\mathcal{Q}_{\mathcal{A}} : \text{Sym}(r) \rightarrow \mathbb{R}$. In the limit $n \rightarrow \infty$ the maximal term \mathcal{Q}_{\max} is dominant in $\beta_r^{\text{reg}}(\mathcal{A})$ and therefore the only relevant term. However, finding \mathcal{Q}_{\max} is not easy in general. In the case of the qubit depolarizing channel we proved that $\mathcal{Q}_{\max} = \max\{2, \mathcal{Q}_{\Delta_\lambda}((1 \dots r))\}$ and also that $\bar{\beta}_r(\Delta_\lambda) = \bar{\mathcal{S}}_r^{\text{reg}}(\Delta_\lambda)$ for some λ . For all $r \in \mathbb{N}$ and $\lambda \leq 1/3$ the regularized output entropy becomes 1. Because the typical high-dimensional random state is highly entangled our result for $\lambda \leq 1/3$ agrees with the notion that the tensor product of an entanglement breaking channel “chops up” these highly entangled states and produces maximally mixed states with almost certain probability.

For future work it would be interesting to also study channels other than the depolarizing channel, especially channels that are not known to be additive. Also, it might be of interest to study quantities similar to $\bar{\mathcal{S}}_r^{\text{reg}}$ with the same general procedure, for example $\mathbb{E}[\text{Tr } \mathcal{A} \otimes \mathbb{1}(|\phi\rangle\langle\phi|)^r]$. It would be insightful to gain a better understanding of the typical output and input states by finding explicit examples that conform with the average output.

6 Acknowledgements

We thank the reviewer for important corrections and improvements which helped the quality and presentation of the paper.

A \mathcal{Q} sum diagrams

We calculate an example of the subsums appearing in the proof of Lemma 12. Let $\alpha = (123)(45)$, $\beta = (143)(25)$ and $\gamma = (24)$. With this choice the correspondence of diagrams consists of switching tiles 2 and 4. The choice of indices

x_1	x_2	x_3	x_4	x_5
0	1	1	0	0

gives the diagrams and totals shown in Table 2. The contribution to $\mathcal{Q}(\alpha)$ dominates in both factors as expected because α is non-overlapping.

References

- [1] G. Smith, “Quantum channel capacities,” arxiv:1007.2855v1, 2010.
- [2] G. Amosov, A. S. Holevo, and R. F. Werner, “On some additivity problems in qit,” *Problems in information transmission*, vol. 36, pp. 305–313, 2000.
- [3] K. Matsumoto, T. Shiono, and A. Winter, “Remarks on additivity of the holevo channel capacity and of the entanglement of formation,” *Commun. Math. Phys.*, vol. 246, pp. 427–442, 2004.

	$\lambda^2\mu^3 + \kappa^2\nu^3$	\leftrightarrow		$\lambda^2\mu\nu^2 + \kappa^2\nu\mu^2$
	$\lambda^2\mu\nu^2 + \kappa^2\nu\mu^2$	\leftrightarrow		$\lambda^2\mu\nu^2 + \kappa^2\nu\mu^2$
	$\kappa^2\nu\mu^2 + \lambda^2\mu\nu^2$	\leftrightarrow		$\kappa^2\nu\mu^2 + \lambda^2\mu\nu^2$
	$\kappa^2\nu^3 + \lambda^2\mu^3$	\leftrightarrow		$\kappa^2\nu\mu^2 + \lambda^2\mu\nu^2$
Total for $\mathcal{Q}(\alpha)$			Total for $\mathcal{Q}(\beta)$	
$2(\lambda^2\mu + \kappa^2\nu)(\mu^2 + \nu^2)$			$2(\lambda^2\nu + \kappa^2\mu)(\mu\nu + \nu\mu)$	

Table 2: Contributions to the sums $\mathcal{Q}(\alpha)$ and $\mathcal{Q}(\beta)$.

- [4] P. Hayden and A. Winter, “Counterexamples to the maximal p-norm multiplicativity conjecture for all $p \geq 1$,” *Commun. Math. Phys.*, vol. 284, pp. 263–280, 2008.
- [5] M. B. Hastings, “Superadditivity of communication capacity using entangled inputs,” *Nature Physics*, vol. 5, pp. 255–257, 2009.
- [6] M. Ruskai, “Qubit entanglement breaking channels,” *Rev. Math. Phys.*, vol. 15, pp. 643–662, 2003.
- [7] M. Horodecki, P. W. Shor, and M. Ruskai, “General entanglement breaking channels,” *Rev. Math. Phys.*, vol. 15, pp. 629–641, 2003.
- [8] A. S. Holevo, “Quantum coding theorems,” *Russian Math. Surveys*, vol. 53, pp. 1295–1331, 1999.
- [9] M. D. Choi, “Completely positive linear maps on complex matrices,” *Linear Algebra Appl.*, vol. 10, pp. 285–290, 1975.
- [10] C. King, M. Nathanson, and M. B. Ruskai, “Multiplicativity properties of entrywise positive maps,” *Linear Algebra and its Applications*, vol. 404, pp. 367–379, 2005.
- [11] M. Ruskai, S. Szarek, and E. Werner, “An analysis of completely-positive trace-preserving maps on \mathcal{M}_2 ,” *Linear Algebra Appl.*, vol. 347, pp. 159–187, 2002.
- [12] M. Ledoux, “The concentration of measure phenomenon,” *Mathematical Surveys and Monographs*, vol. 89, 2001.
- [13] V. D. Milman and G. Schechtman, “Asymptotic theory of finite dimensional normed spaces,” *Lecture Notes in Mathematics*, vol. 1200, 1986.
- [14] P. Hayden, D. W. Leung, and A. Winter, “Aspects of generic entanglement,” *Commun. Math. Phys.*, vol. 265, pp. 95–117, 2006.

- [15] M. Fukuda and C. King, “Entanglement of random subspaces via the hastings bound,” *J. Math. Phys.*, vol. 51, 2010.
- [16] G. Aubrun, S. Szarek, and E. Werner, “Hastings’ additivity counterexample via dvoretzky’s theorem,” *Commun. Math. Phys.*, 2010.
- [17] D. Weingarten, “Asymptotic behavior of group integrals in the limit of infinite rank,” *J. Math. Phys.*, vol. 19, pp. 999–1001, 1978.
- [18] B. Collins and I. Nechita, “Random quantum channels I: graphical calculus and the bell state phenomenon,” *Commun. Math. Phys.*, vol. 297, pp. 345–370, 2009.
- [19] A. J. Hoffman and H. W. Wielandt, “The variation of the spectrum of a normal matrix,” *Duke Math. J.*, vol. 20, pp. 37–39, 1953.